

January 7, 2022

Submitted via ECF

Magistrate Judge Alex G. Tse
San Francisco Courthouse
Courtroom A - 15th Floor
450 Golden Gate Avenue
San Francisco, CA 94102

Re: Joint Letter Brief re: Google Search Terms
Rodriguez v. Google LLC, Case No. 3:20-cv-04688-RS (N.D. Cal.)

Dear Magistrate Judge Tse:

Plaintiffs and Google LLC (“Google”) submit this joint letter brief regarding Plaintiffs’ request for an order requiring Google to use additional search terms on the first five of 24 custodians’ ESI: Dave Monsees, Steve Ganem, Francis Ma, Edward Weng, and Todd Hansen. The parties have started the process of negotiating search terms to be applied to the 19 new custodians authorized by the Court (ECF No. 184), but in the interest of expeditiousness have opted to raise this issue with the Court now. The parties may file a subsequent letter brief regarding search terms for the 19 new custodians, which may include terms discussed in this letter brief.

Counsel for the parties have exchanged meet-and-confer letters, and met and conferred by videoconference on October 29 and November 4. Ex. A includes Plaintiffs’ proposed order, and Ex. B includes Google’s proposed order.

PLAINTIFFS' STATEMENT

Plaintiffs ask the Court to order Google to use all of the search terms listed in Plaintiffs' Proposed Order (Exhibit A). Google's burden complaints are unfounded, particularly in light of the fact that Plaintiffs have repeatedly offered reasonable compromises while Google has hidden the ball throughout the search-term negotiation process. Prior to sending Plaintiffs its first draft of this letter brief, Google refused to provide a single hit count for any search term, effectively asking Plaintiffs to negotiate against themselves. Then, Google's first draft of this brief primarily complained that Plaintiffs' proposal would have required Google to review over a million additional documents. Plaintiffs responded by proposing yet another compromise, which cut 500,000 documents from the hit counts. Yet Google refuses to budge. Google's portion below also tells a story that never happened. There were no "repeated[] compromise[s]" from Google. Rather, Google dictated the initial search terms it would run, and Plaintiffs, consistent with the Court's instructions for selecting custodians, decided to first review Google's initial productions before asking for additional custodians and search terms. Plaintiffs now seek the following search terms, organized by category for ease of discussion.

I. WAA

Google should search for "WAA" and all variations thereof across all custodians. WAA is central to this case. Google employees' discussions about WAA and Google's disclosures are highly relevant, particularly insofar as they reveal Google's awareness of its false representations. These terms should especially be run for Mr. Monsees, who played a central role in crafting the disclosures at the heart of Plaintiffs' claims. *See, e.g.*, GOOG-RDGZ-00024690 ("I have discussed rewrites of all the UDC help center articles (including WAA) to create a more consistent structure to explain the kind of 'if it's on'; vs. 'if it's off' points"); GOOG-RDGZ-00024493 (listing Mr. Monsees as the "Product review" for the "See & control your Web & App Activity" Help Center webpage that Chief Judge Seeborg relied on in denying in part Google's initial motion to dismiss).

Google tries to downplay the importance of "WAA" as a search term, suggesting that Plaintiffs are merely "suspicious" that documents are missing. But Plaintiffs know that Google's chosen terms are inadequate. They omit key emails like one in which Mr. Monsees admitted that WAA "confuses users." GOOG-RDGZ-00015004. That email was only produced because of this Court's Order for Google to search its Arizona productions. *See* Dkt. No. 85. Google's reliance on *McAfee, Inc.*, 2014 WL 5280966 is misplaced, particularly because the quote Google cites came from the parties' ESI Order. In any event, *McAfee* is inapposite because just three proposed search terms regarding "product names" would have generated "almost 500,000 results"—far more than the measly 27,180 hits for *WAA*. Ex. C at 4. The more apt case law comparison is the *Brown v. Google* case regarding Google's collection of private browsing data, where Google was ordered to run "Incognito" for all custodians. 20-cv-03664 at Dkt. 148. Finally, Google's suggestion that "WAA" should be combined with product names like GA4F is disingenuous because it contradicts Google's argument that WAA is "completely unrelated" to GA for Firebase. Dkt. 62 at 1.

Plaintiffs here briefly explain some of the WAA terms. "sWAA" (sometimes "supplemental Web & App Activity" or "additional Web & App Activity") is a subset of WAA specific to data derived from visits to third-party apps that "use Google services." Dkt. 138, TAC

¶ 247. “UDC” / “User Data Controls” are other ways that Google employees refer to WAA, among other Google settings. *See e.g.*, GOOG-RDGZ-00024699. The WAA button itself is found on Google’s “Activity controls” webpage. TAC ¶ 247. And “SMH, “my history,” and myhistory are “equivalent to” WAA. GOOG-RDGZ-00014889.

II. Google’s Representations to Users:

These search terms are necessary to ensure that Google produces documents relating to the disclosures on which Plaintiffs ground their claims. *See* TAC ¶¶ 65-79, 245-58. These terms are narrowly tailored to capture key language within the disclosures. *E.g.*, “must be on.” “Privacy Policy” and “Help Center” are crucial because that is where many of these key Google disclosures are located. *See* Dkt. 109 at 5, 8 (relying on Plaintiffs’ interpretation of the Privacy Policy and the Help Center’s WAA disclosures). Google is taking the position that its disclosures are clear and accurate; these search terms will yield documents which show that Google’s employees disagree.

III. The Relevant Google Products and Services

Plaintiffs initially asked that Google individually run all variations of GA4F, AdMob, and FCM for each custodian. As a compromise, Plaintiffs later proposed more targeted search terms that combine each product with other relevant products (i.e., [AdMob AND (GA4F OR “cloud messaging” OR . . .)]). Plaintiffs even designed different terms for each custodian. Plaintiffs now stand by their reasonable compromise proposals, including because the hit counts show that Plaintiffs struck the right balance. *See* Ex. C at 10-12.

Plaintiffs here briefly explain why each of the limiters in their proposals is relevant. “Google Mobile Service” or GMS is what Google uses to collect and store the intercepted app-interaction data (on Android devices). Even Google has repeatedly relied on Plaintiffs’ GMS allegations in their motions to dismiss. Dkts. 115 at 18, 139 at 4. “Clearcut” is how Google logs data within GMS. GOOG-RDGZ-00038319. “Scion,” “Google Mobile Platform,” and “Gold” are associated with GA for Firebase. GOOG-RDGZ-00056947; GOOG-RDGZ-00057081. And terms like “event,” “interact,” “activity,” “measure,” “perform,” “conversion,” and “audience” help target the searches to the specific app-interaction data collected and the ways in which it is used. Plaintiffs are also entitled to discovery on Real-time bidding (“RTB”) because Google has admitted that “data collected through GA for Firebase can be used for RTB [real-time bidding] purposes.” Dkt. 145 at 5. Plaintiffs also seek a targeted search term with “App Indexing” to understand whether WAA affects App Indexing’s collection of user data. TAC ¶ 53 n.14.

IV. Google’s Consent Defense:

These terms will generate documents that refute Google’s consent defense. Dkts. 62 at 8, 139 at 10. “CFS is a backend service that provides the right consent texts to opt-in users to some of Google’s key privacy settings (like the UDC bits).” GOOG-RDGZ-00018757. Na■■■■ (later N2 and N3) is a “company-wide user trust OKR to give signed-in users better control and transparency for their ads experience, requiring a global consent effort.” GOOG-RDGZ-00019099. “Consent Bump” “is a critical piece of the Na■■■■ 2.0 effort focused on educating users about a change to how Google collects and processes user data.” GOOG-RDGZ-00018662. Google suggests below

that its consent defense is limited to “third-party app developers’ disclosures” to users. This is not true. Google also argues that its own disclosures are clear. Dkt. 62 at 6.

V. Google’s Storage and Use of Data:

As explained in Plaintiffs’ supplemental letter brief regarding their request for additional Google custodians (Dkt. 163), the parties have a factual dispute about how Google stores and uses the data it impermissibly collects. While Plaintiffs allege that the data “is not anonymized,” TAC ¶ 126, Google argues that the data is “pseudonymous, i.e., not tied to any particular user,” Dkt. 155 at 5. Google should be required to run search terms that will yield documents relevant to this dispute. Zwieback, Biscotti, K█████, GAIA, Footprints, Plogs, GWS, Sawmill and BigQuery are relevant to how Google logs and tracks the data by a user’s device. *See* GOOG-RDGZ-00046758 (discussing “change from Zwieback to Gaia-keyed logs” for WAA-off data, including within “plogs” and “GWS”); GOOG-RDGZ-00036800 (discussing using GAIA (Google-account identifier) to track WAA-off data); GOOG-RDGZ-00042121 (discussing using Biscotti cookie identifiers to log WAA-off data); GOOG-RDGZ-00045173 (discussing how WAA-off data is stored in “K█████”); GOOG-RDGZ-00019098 (discussing storage of “app activity data” within Footprints); GOOG-RDGZ-00028267 (describing “BigQuery” as a “warehouse” that “facilitates” “cross-platform tracking”). “Firelog,” “Firedata,” and “Firestore” are a “new server” being built as a “clearcut replacement” to store the data collected by GMS. GOOG-RDGZ-00059505. The remaining terms are designed to capture relevant ids that Google uses to track the data.

GOOGLE’S STATEMENT

Plaintiffs already have the documents they need to understand the accused products because Google’s comprehensive searches have yielded them. Plaintiffs’ new request is wildly overbroad: applied to the first five custodians, Plaintiffs would have Google review almost *600,000 documents*, which will take ~17,000 hours at a cost of over \$1 million. *See* Ex. C; Santacana Decl. ¶¶ 3–4. Plaintiffs’ compound search strings—which comprise 390 unique searches—hide a multitude of overbroad searches with “AND” and “OR” logic. And Plaintiffs’ single-word searches are equally problematic; terms like “K█████” and “GAIA” each return ~110,000 documents, yet neither is limited to relevant issues. What’s more, these figures don’t account for the 19 recently added custodians, which Plaintiffs are already arguing should be subject to all search terms used for these five custodians; extrapolating from these numbers, if their wish were made reality, that would add millions more documents and several million dollars of review costs.

Plaintiffs’ approach runs counter to this District’s best practices, which require cooperation to achieve a reasonable scope of review. N.D. Cal. ESI Guidelines at § 1.02; ESI Order, ECF No. 73 at ¶ 5(d). Google repeatedly compromised over nine months, agreeing to review ESI from over 200 unique searches *proposed by Plaintiffs*. Plaintiffs claim that this Court Ordered them to lie in wait until now, and that Google “dictated” the search terms it would run. But to date, Google has never proposed a single search term. Every one of the terms Google has run was proposed by Plaintiffs in successive rounds of proposals long after this Court’s first Order on custodians. Google accepted the vast majority of those terms and ran them or variations of them. Plaintiffs silently accepted the results of those searches (including the setting of a substantial completion deadline). Plaintiffs now circle back to demand 100% of their original requested search terms as well as hundreds more. Google’s alternative proposal (below) is reasonable and proportionate, and the Court should adopt it instead.

Google's work so far searching for relevant documents justifies denial of the motion. **Exhibit C** lists Google's searches for each topic identified by Plaintiffs and compares Plaintiffs' current demand for additional searches and review of 600,000 more documents (for five custodians) (right columns) with Google's compromise proposal to review ~33,000 more documents (left columns). Plaintiffs purport to identify search terms designed to capture highly relevant documents, but the numbers don't lie. Plaintiffs' terms are overbroad and ignore basic tenets of search term methodology, as well as Rule 26(b)(1)'s proportionality standard, which considers, *inter alia*, "the importance of the discovery in resolving the issues, and whether the burden or expense . . . outweighs its likely benefit." *See also Elgindy v. AGA Service Co.*, 2021 WL 5083761, at *2 (N.D. Cal. Nov. 2, 2021). A low hit count alone cannot make WAA-related terms appropriate if Plaintiffs already understand how WAA works, nor can WAA's relevance *per se* warrant a large-scale review. An overbroad but marginally relevant term (e.g., "help center") is similarly inappropriate because it would require Google to manually filter out a vast number of irrelevant documents (e.g., Google Cloud Help Center pages). Considering Google's efforts thus far, the remaining terms at issue are disproportionate, unlikely to resolve disputed issues, or otherwise unduly burdensome. Google briefly addresses each category below.

WAA Terms (Ex. C at 2–4). Plaintiffs propose WAA terms—without any limiters—for David Monsees, Google's WAA Product Manager. This is disproportionate to the needs of the case because Mr. Monsees's job is WAA-focused, so running WAA terms without limitation would hit on nearly all of his ESI. Instead, before custodial review commenced, Google produced Mr. Monsees's documents from the Arizona litigation per this Court's Order. Google also ran targeted WAA terms for Mr. Monsees by limiting stand-alone WAA terms with others proposed by Plaintiffs. *See* Ex. C at 2–3. Searching WAA without limiters now would be unreasonable. WAA covers more than GA for Firebase; it also relates to, e.g., online web search history and YouTube history, neither of which are implicated by the TAC. Because Google would have to manually filter out those irrelevant documents, the term is disproportionate. *See, e.g., TVIIM, LLC v. McAfee*, 2014 WL 5280966, at *2 (N.D. Cal. Oct. 15, 2014) (ESI order noting "[i]ndiscriminate terms, such as . . . product name[s], are inappropriate [absent] narrowing . . . criteria"). Plaintiffs know this: at the beginning of the search-term selection process, they agreed to "use . . . certain [WAA]-related terms for individuals *not* responsible for [WAA]" and Firebase terms for the WAA custodian. Pltfs' 2/24/21 Letter at 5. They chose not to move to compel on this issue then, and Google relied on that agreement. Plaintiffs' eleventh-hour about-face seeking standalone WAA terms from the WAA custodian is baseless; nothing has changed. Indeed, when pressed, Plaintiffs were unable to articulate any aspect of WAA that they don't now understand; they simply stated that they'd like to use these terms "to be complete." This, by itself, is fatal: "completeness" is the opposite of what Rule 26 counsels. Plaintiffs' general "suspicion that all responsive documents have not been produced . . . [is] insufficient to warrant an order compelling" review. *See Swanson v. ALZA Corp.*, 2013 WL 5538908, at *3 (N.D. Cal. Oct. 7, 2013).

Other Products (Ex. C at 10–14). Plaintiffs' request that Google employ search strings related to GA for Firebase, AdMob, and FCM on the custodians respectively responsible for those products is inappropriate for the same reasons: it will result in a vastly overbroad set of documents unconnected to the case. That is why Google offered a compromise with reasonable limiters. *See* Ex. C at 10–14, col. 2. But Plaintiffs demand that Google employ search strings with terms as broad as "Firebase" (which comprises 19 products, 17 of which are not at issue in this case) and "event" and "perform*". Those proposals (*see* Ex. A at 3–4) are replete with the connector "OR,"

meaning they will hit on irrelevant documents that could, for example, simply mention one product in passing without any relevance while imposing a crushing review burden. As for GMS, publicly available information confirms that it is colossally broader than this case: “GMS[] is a collection of Google applications,” such as Gmail and YouTube, and hundreds of APIs like ones on web fonts. *See, e.g.,* android.com/gms. So is FCM, which has essentially nothing to do with this case except for a single integration that exists between it and Google Analytics.

Google’s Representations (Ex. C at 6–7). Plaintiffs assert that their representation-related terms are “necessary to ensure that Google produces documents relating to [] disclosures.” But they do not explain why the terms Google already used aren’t enough. Plaintiffs point to parts of the TAC that concern disclosures *related to WAA* (*see* TAC ¶¶ 65–79, 245–58). But those documents were identified with the WAA-related terms Google already used. Plaintiffs also contend their terms are “narrowly tailored to capture key language,” but that claim is belied by the lack of limitations on terms like “help center.” *See* Ex. C at 7, row 13. Google’s “Help Center” covers thousands of pages for hundreds of products; anything relevant was captured by product-specific terms. This search hits on over 16,000 documents.

Google’s Consent Defense (Ex. C at 6–7). Plaintiffs’ new terms they deem to be “relevant to Google’s consent defense” are improper for at least two reasons. *First*, Google has already run numerous search strings that Plaintiffs proposed concerning that defense. *See* Ex. C at 6. The newly proposed ones are overbroad because they seek any document concerning a discussion of users’ understanding of *anything*. *See* Ex. C at 6, row 10. Google has many users who use many different products and services; the terms Plaintiffs propose on Google’s consent defense are in no way tailored to the claims of this case. Likewise, the term proposed at row 17 (“Na OR N2 OR N3”) results in tens of thousands of hits relating to a project that has little to nothing to do with this case. This term is also not limited by any tailoring that relates to the claims here. And if they were, they would simply mirror the terms Google already ran. *Second*, even assuming these terms could be appropriately limited, that would be unwarranted due to Plaintiffs’ unjustified delay. Plaintiffs *never* proposed many of the new terms (e.g., NAC, N2, etc.) in the dozens of discovery conferences and hundreds of emails and letters exchanged over the last ten months. But Plaintiffs have had *every* document they cite since *before* Google began its custodial review. This request is little more than an effort to blow up the scope of discovery unnecessarily, and it upends the collaborative search-term negotiation procedure encouraged by courts. *E.g., Baranco v. Ford Motor Co.*, 2018 WL 9869540, at *1 (N.D. Cal. Apr. 10, 2018).

Storage and Use of Data (Ex. C at 8–9). Plaintiffs contend that Google should search broad, single-word terms they claim are “relevant to how Google logs and tracks [] data by a user’s device” because they relate to whether Google stores data pseudonymously. But nothing in the documents Plaintiffs cite describe logging of non-pseudonymous, *i.e., personally-identifiable information*, when WAA is off, nor are the terms Plaintiffs suggest limited to logging of data received through GA for Firebase (the relevant data here). In addition to being irrelevant, these terms are so “generalized” they are overly broad on their face and would thus “pose an undue burden on defendants” regardless of the time review would take. *See Brunston v. Bayer Healthcare Pharms.*, 2014 WL 12587032, at *4 (C.D. Cal. May 16, 2014). And, Google has already provided a 9-page interrogatory response detailing such usage and storage and Plaintiffs have sought Rule 30(b)(6) testimony on that topic.

Respectfully,

WILLKIE FARR & GALLAGHER, LLP

SUSMAN GODFREY LLP

By: /s/ Eduardo E. Santacana

Eduardo E. Santacana (SBN: 281668)
esantacana@willkie.com
Benedict Y. Hur (SBN: 224018)
bhur@willkie.com
Simona Agnolucci (SBN: 246943)
sagnolucci@willkie.com
Lori Arakaki (SBN: 315119)
larakaki@willkie.com
One Front Street, 34th Floor
San Francisco, CA 94111
Telephone: (415) 858-7400
Facsimile: (415) 858-7599

Attorneys for Defendant Google LLC

/s/ Amanda Bonn

Amanda Bonn (CA Bar No. 270891)
Mark C. Mao (CA Bar No. 236165)
mmao@bsfllp.com
Beko Reblitz-Richardson (CA Bar No.
238027)
brichardson@bsfllp.com
BOIES SCHILLER FLEXNER LLP
44 Montgomery Street, 41st Floor
San Francisco, CA 94104
Telephone: (415) 293 6858
Facsimile (415) 999 9695

James W. Lee (*pro hac vice*)
jlee@bsfllp.com
Rossana Baeza (*pro hac vice*)
rbaeza@bsfllp.com
100 SE 2nd Street, Suite 2800
Miami, FL 33130
Telephone: (305) 539-8400
Facsimile: (305) 539-1304

William Christopher Carmody (*pro hac vice*)
bcarmody@susmangodfrey.com
Shawn J. Rabin (*pro hac vice*)
srabin@susmangodfrey.com
Steven Shepard (*pro hac vice*)
sshepard@susmangodfrey.com
SUSMAN GODFREY L.L.P.
1301 Avenue of the Americas, 32nd Floor
New York, NY 10019
Telephone: (212) 336-8330

Amanda Bonn (CA Bar No. 270891)
abonn@susmangodfrey.com
SUSMAN GODFREY L.L.P.
1900 Avenue of the Stars, Suite 1400

Los Angeles, CA 90067
Telephone: (310) 789-3100

John A. Yanchunis (*pro hac vice*)
jyanchunis@forthepeople.com
Ryan J. McGee (*pro hac vice*)
rmcgee@forthepeople.com
MORGAN & MORGAN, P.A.
201 N Franklin Street, 7th Floor
Tampa, FL 33602
Telephone: (813) 223-5505
Facsimile: (813) 222-4736

Attorneys for Plaintiffs

ATTESTATION

I, Amanda Bonn, hereby attest, pursuant to N.D. Cal. Civil L.R. 5-1, that the concurrence to the filing of this document has been obtained from each signatory hereto.

DATED: January 7, 2022

By: /s/ Amanda Bonn